Proceedings of the AMS Centennial Symposium
August 8–12, 1988

A Century of Lie Theory

ROGER HOWE

The origins of a subject are frequently difficult to trace. The extent to which precursor fields and early investigations, later perceived to have anticipated the emergence of a field or to fit naturally into it, should be annexed to the field can be a matter for vigorous debate. In the case of Lie theory, Sophus Lie was already studying "continuous, finite groups of transformations" in the 1870s, and one could even make a case for including Euclidean geometry as part of Lie theory **[Crtn7]**. However, in 1888 the first volume of *Theorie der Transformationsgruppen* by "S. Lie unter Mitwirkung von Dr. F. Engel" **[LiEn1]** was published by Teubner in Leipzig; and also Killing's classification **[Kill]** of complex semisimple Lie algebras appeared in Mathematische Annalen. These events are the basis for the title of this article.

My assignment, roughly, is to report on the development of Lie theory over the past 100 years, and to extrapolate it into the future. To do this in a uniform, systematic way is, for me and I suspect for anyone, impossible. So this account will be frankly idiosyncratic; I make here a blanket apology to the many investigators whose interesting results will be slighted or ignored completely; or maybe worse, treated clumsily. All I can offer by way of consolation is the remark that it has happened to me too. Similarly, although the bibliography is extensive, it is not at all comprehensive. References are only intended to provide the reader with representative sources of further information. Again I offer apologies to the many authors who will find that I have neglected to mention relevant work of theirs.

1. The first example of a Lie group is Euclidean space $\mathbf{R}^n$ with vector addition as the group operation, but it is too simple-minded a group to reveal the essential features of Lie theory. Almost as well known, and much more interesting structurally, is $\mathrm{GL}_n(\mathbf{R})$, the group of real invertible $n \times n$ matrices, with matrix multiplication as the group operation. It serves as the basic template for Lie theory in the following sense: any subgroup of $\mathrm{GL}_n(\mathbf{R})$ which is closed (with respect to the standard topology on $n \times n$ matrices) is

1991 *Mathematics Subject Classification.* Primary 22-02.

0-8218-0167-8 \$1.00 + \$.25 per page

a Lie group; and conversely, modulo some relatively subtle caveats ignorable at this point (see §2.5), any Lie group is realizable as a closed subgroup of $\mathrm{GL}_n(\mathbf{R})$ for some n (cf. **[Hoch]**). Of course, $\mathrm{GL}_n(\mathbf{R})$ contains discrete, even finite, subgroups, but Lie theory, in its most basic form, ignores these. The subgroups of $\mathrm{GL}_n(\mathbf{R})$ which are the immediate subjects of Lie theory are the ones which are the opposite of discrete: the connected ones. The first miracle of Lie theory is that the extremely weak topological hypothesis—closed and connected, when combined with the algebraic condition—subgroup, yields a subset which is a smooth (even analytic) surface (i.e., submanifold). If one then looks at the tangent space to this surface at the identity matrix, one finds it is endowed with a certain algebraic structure, the Lie bracket (which as an operation on matrices is simply commutator). This is the Lie algebra. The second miracle of Lie theory is that, except for the caveats ignored above, this Lie algebra, a vector space with a bilinear nonassociative product, completely determines the group from which it comes.

Without discussing in detail yet the foundational results of Lie theory, we can observe that its essential feature seems to be the enrichment of the algebraic notion of group by the topological notion of continuity: a Lie group is an object which carries in a compatible way the structure of group and of differentiable manifold (in fact analytic manifold; by Hilbert's 5th Problem, finally solved in the early 1950s **[Glea, MoZi1, Yama1, Yama2, MoZi2, Kapl1]**, it is enough to require a Lie group to be locally homeomorphic to Euclidean space—no smoothness need be explicitly assumed). Thus continuity, indeed smoothness, seems to be a *sine qua non* of the theory. Therefore, it is interesting to observe that Lie theory has intimate and fruitful interactions with the theory of discrete, in particular finite, groups.

1.1. An important aspect of the connection can be illustrated by a careful study of the bread-and-butter topic of elementary linear algebra, Gaussian elimination. Let A be an $n \times n$ matrix with entries $\{a_{ij} : 1 \leq i, j \leq n\}$, and let $z = (z_1, z_2, \dots, z_n)^{\mathrm{T}}$ be a column vector of length n. Consider the system of linear equations

$$Ax = z, \tag{1.1.1}$$

from which we want to solve another column vector of length n for x. We will assume A is invertible, so that the solution x exists and is unique. Gaussian elimination is a standard method for solving system (1.1.1). We will discuss it in its naive form, untouched by worries about round-off error.

Writing the system (1.1.1) out long-hand we obtain:

$$\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{in}x_n = z_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = z_2 \\ \vdots \qquad\qquad\qquad \vdots \quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = z_n \end{array} \tag{1.1.2}$$

Suppose $a_{11} \neq 0$. Then for $k = 2, 3, \ldots, n$, we can subtract a_{k1}/a_{11} times the first equation from the kth equation to arrive at an equivalent system:

$$\begin{aligned} a'_{11}x_1 + a'_{12}x_2 + a'_{13}x_3 + \cdots + a'_{1n}x_n &= z'_1 \\ a'_{22}x_2 + a'_{23}x_3 + \cdots + a'_{2n}x_n &= z'_2 \\ \vdots \qquad\quad \vdots \qquad\quad &\ \ \vdots \\ a'_{n2}x_2 + a'_{n3}x_3 + \cdots + a'_{nn}x_n &= z'_n \end{aligned} \tag{1.1.3}$$

where $a'_{1j} = a_{1j}$, $z'_1 = z_1$, and $a'_{kj} = a_{kj} - a_{k1}a_{1j}a_{11}^{-1}$, $z'_k = z_k - a_{k1}z_1a_{11}^{-1}$, $k \geq 2$.

Since the last $n-1$ equations of system (1.1.3) involve only the $n-1$ unknowns $x_2, x_3, \ldots, x_n$, we evidently have a recursive procedure for solving the system (1.1.1)–(1.1.2). Providing that $a'_{22} \neq 0$, we may subtract a'_{k2}/a'_{22} times the second equation in system (1.1.3) from the third through nth equations to obtain a third equivalent system with a subsystem of $n-2$ equations in $n-2$ unknowns. And soon, after $n-1$ steps, we will arrive at a triangular system:

$$\begin{aligned} b_{11}x_1 + b_{12}x_2 + \cdots + b_{1n}x_n &= y_1 \\ b_{22}x_2 + \cdots + b_{2n}x_n &= y_2 \\ b_{33}x_3 + \cdots + b_{3n}x_n &= y_3 \\ &\ \ \vdots \\ b_{nn}x_n &= y_n \end{aligned} \tag{1.1.4}$$

This system can of course be solved by "back-substitution." Of the several slight variants of this procedure, we select the following. First, divide each equation by its leading coefficient to obtain:

$$\begin{aligned} x_1 + b'_{12}x_2 + b'_{13}x_3 + \cdots + b'_{1n}x_n &= y'_1 \\ x_2 + b'_{23}x_3 + \cdots + b'_{2n}x_n &= y'_2 \\ &\ \ \vdots \\ x_n &= y'_n \end{aligned} \tag{1.1.5}$$

where $b'_{ij} = b_{ii}^{-1}b_{ij}$, $y'_i = b_{ii}^{-1}y_i$. Now observe we have already solved for x_n:

$$x_n = y'_n. \tag{1.1.6a}$$

We can therefore compute x_{n-1} by the simple recipe

$$x_{n-1} = y'_{n-1} - b'_{n-1n}x_n, \tag{1.1.6b}$$

and so on. If we know $x_n, x_{n-1}, \ldots, x_{i+1}$, then we compute x_i by the formula

$$x_i = y'_i - \sum_{j=i+1}^{n} b'_{ij}x_j. \tag{1.1.6c}$$

Thus, under some mild assumptions about the nonvanishing of certain numbers, we have a systematic procedure for performing matrix inversion by means of ordinary (i.e., scalar) arithmetic.

Let us formulate this procedure in terms of matrix manipulations. Let L_1 be the matrix

$$L_1 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -a_{11}^{-1}a_{12} & 1 & 0 & \cdots & 0 \\ -a_{11}^{-1}a_{13} & 0 & 1 & \cdots & 0 \\ & & & \ddots & \vdots \\ & & & & 0 \\ -a_{11}^{-1}a_{1n} & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{1.1.7}$$

Then in terms of matrices, the passage from the system (1.1.2) to the system (1.1.3) amounts to the multiplication of both sides of equation (1.1.1) by L_1: equation (1.1.1) is the matrix version of (1.1.2) and the matrix version of (1.1.3) is

$$L_1Ax = L_1y. \tag{1.1.8}$$

Similarly, the second stage of the procedure is equivalent to multiplying the system (1.1.8) by the matrix

$$L_2 = \begin{bmatrix} 1 & 0 & 0 & \cdots & & 0 \\ 0 & 1 & 0 & \cdots & & 0 \\ 0 & (-a'_{22})^{-1}a'_{21} & 1 & 0 & \cdots & 0 \\ 0 & (-a'_{22})^{-1}a'_{31} & 0 & 1 & & 0 \\ \vdots & \vdots & & & & \vdots \\ 0 & 0 & & & & 0 \\ 0 & (-a'_{22})^{-1}a'_{n1} & 0 & 0 & \cdots & 1 \end{bmatrix}. \tag{1.1.9}$$

Of the matrices L_i, we may observe

(1.1.10)

(i) Each L_i has all entries zero above the diagonal; that is, L_i is lower triangular.

(ii) Additionally each L_i has all its diagonal entries equal to 1. Since for a lower triangular matrix the diagonal entries equal the eigenvalues, this is the same as saying all the eigenvalues of L_i are equal to 1. A matrix with all eigenvalues equal to 1 is called *unipotent*.

Thus the matrices L_i are unipotent lower triangular matrices. Let us denote the set of all unipotent lower triangular matrices by $\overline{\mathcal{U}}$. We can easily check that

(1.1.11)
(i) The product of two matrices in $\overline{\mathscr{U}}$ is also in $\overline{\mathscr{U}}$.
(ii) The inverse of a matrix in $\overline{\mathscr{U}}$ is in $\overline{\mathscr{U}}$ (the analogous fact for unipotent upper triangular matrices is implicit in equations (1.1.6a)–(1.1.6c)). Thus $\overline{\mathscr{U}}$ is a group. In particular, the successive multiplications of our system (1.1.1) by the L_i may be achieved by multiplication by a single appropriate element of $\overline{\mathscr{U}}$.

The result of these modifications by lower triangular unipotent matrices is the system (1.1.4) which, in contrast to the L_i, is the system corresponding to an *upper* triangular matrix. Thus passage from (1.1.2) to (1.1.4) is expressed in matrix terms by the equation

$$\tilde{L}A = B, \tag{1.1.12}$$

where $\tilde{L}$ is a unipotent lower triangular matrix and B is an upper triangular matrix. Further, the passage from (1.1.4) to (1.1.5) corresponds to a matrix factorization

$$B = DU, \tag{1.1.13}$$

where D is a diagonal matrix and U is a unipotent upper triangular matrix. Combining (1.1.12) and (1.1.13) gives us

$$\tilde{L}A = DU \tag{1.1.14}$$

which we can convert to a factorization

$$A = LDU, \tag{1.1.15}$$

where L $(= \tilde{L}^{-1})$ is lower triangular unipotent, D is diagonal, and U is upper triangular unipotent.

Thus, Gaussian elimination produces, and is equivalent to, the factorization (1.1.15) of a "generic" matrix A into a product of upper and lower triangular unipotent matrices and a diagonal matrix. This equivalence is well known and can be found in elementary textbooks **[Hill, Stra]**.

1.2. However, as we have noted, this procedure will not always work.[1] We can ask what to do when it does not. One can observe that it is always possible to permute the equations so that, after rearrangement, the desired diagonal coefficient is nonzero, and the elimination can proceed. This provides an algorithm that will always work, so elementary texts usually stop their discussion with this, or a similar remark.

However, it is interesting to see what one can do in as systematic a fashion as possible. Let us look again at the system (1.1.1) or (1.1.2), admitting the possibility that $a_{11} = 0$. Then we may search down the first column until we find a nonzero coefficient. (There must be one if A is nonsingular.) Suppose the first row with a nonzero first entry is row i_1. Then we may add a multiple of the i_1th row to the rows below to make zeros of all entries of the first column, except for $a_{i_1 1}$. To describe this precisely, let E_{ij} denote the "matrix unit" which has 1 in the ith row, jth column, and zeros everywhere else. Then the process of eliminating all but one entry in the first column amounts to multiplying A on the left by the matrix

$$L_1 = I - \sum_{j \geq i_1} a_{i_1 1}^{-1} a_{j1} E_{ji_1}. \tag{1.2.1}$$

Here I is the identity matrix. Because i_1 was the index of the first row to have a nonzero entry, this matrix will be unipotent lower triangular. Also, note that if $a_{11} \neq 0$, then $i_1 = 1$, and the matrices L_1 of (1.1.7) and (1.2.1) coincide.

So now we have a matrix

$$A' = L_1 A \tag{1.2.2}$$

which has only one nonzero entry in the first column, in the i_1th row. Look at the second column. Choose the index i_2 so that

(1.2.3)
(i) $i_2 \neq i_1$,
(ii) $a'_{i_2 2} \neq 0$,
(iii) i_2 is as small as possible subject to (i) and (ii).

With i_2 so chosen, we can eliminate all entries in the second column below row i_2 by multiplying by the matrix

$$L_2 = I - \sum_{j \geq i_2} a'^{-1}_{i_2 2} a'_{j2} E_{ji_2}. \tag{1.2.4}$$

The matrix L_2 is unipotent lower triangular. The resulting product

$$A'' = L_2 A' \tag{1.2.5}$$

has the properties

(1.2.6)
(i) Only row i_1 has a nonzero entry in the first column; only rows i_1 and i_2 have nonzero entries in the first and second columns.
(ii) In the second column, all rows below row i_2 have a zero.

Note that, with regard to condition (1.2.6)(ii), we should distinguish two cases: if $i_1 > i_2$, then (1.2.6)(ii) says that $a''_{i_1 2} = 0$, so the second column of A'' will have only one nonzero entry, viz. $a_{i_2 2}$; but if $i_1 < i_2$, then row i_1 of A' passes unchanged to A'', and it may happen that $a''_{i_1 2} \neq 0$.

It should now be evident that we can continue this process of multiplying A by unipotent lower triangular matrices until we produce

$$\widetilde{B} = \widetilde{L} A \tag{1.2.7}$$

with the properties:

(1.2.8)
(i) For each j, $1 \le j \le n$, the matrix $\widetilde{B}$ has exactly j rows with nonzero entries in the first j columns.
(ii) If i_j is the row which has its first nonzero entry in the jth column, then $b_{kj} = 0$ if $k > i_j$.

These properties are the analogs for general i of properties (1.2.6) for the case $i = 2$.

Given a matrix B satisfying conditions (1.2.8) we can produce from it an upper triangular matrix B simply by permuting its rows: we move row i_1 to row 1, row i_2 to row 2, and so forth. This also amounts to a matrix multiplication:

$$B = \widetilde{P}\widetilde{B}, \tag{1.2.9}$$

where $\widetilde{P}$ is a permutation matrix—a matrix with all entries zero except for one 1 in each row and column, whose effect on a column n-vector is simply to permute its entries. (Precisely, $\widetilde{P}$ will take the i_jth entry to the jth entry.)

Finally, we can factor the upper triangular matrix B as in (1.1.13). Combining (1.1.13), (1.2.7), and (1.2.9) gives the following result.

THEOREM 1.2.10. *Given an arbitrary invertible $n \times n$ matrix A there is a factorization*

$$A = LPDU, \tag{1.2.11}$$

where

(i) *L is unipotent lower triangular,*
(ii) *P is a permutation matrix,*
(iii) *D is diagonal, and*
(iv) *U is unipotent upper triangular.*

The factors P and D are uniquely determined. Further, U can be made to satisfy the following condition:

(1.2.12) Let p be the permutation of $\{1, 2, \ldots, n\}$ corresponding to the permutation matrix P. If $k < l$, but $p(k) > p(l)$, then u_{kl} (the (k, l)th entry of U) is zero.

If U satisfies condition (1.2.12), then L and U are also uniquely determined.

REMARKS. (a) The factor L in (1.2.11) is related to $\widetilde{L}$ in (1.2.7) by $L = \widetilde{L}^{-1}$. The factor P in (1.2.11) is related to $\widetilde{P}$ in (1.2.9) by $P = \widetilde{P}^{-1}$.

(b) Condition (1.2.12) is just a translation of property (1.2.8)(ii) because, in the notation of (1.2.8), the permutation p will send j to i_j. If $k < l$ and $i_k > i_l$, condition (1.2.8)(ii) says $\tilde{b}_{i_k l} = 0$; but $\tilde{b}_{i_k l} = b_{kl} = b_{kk}u_{kl}$. (Here the $\tilde{b}_{ij}$ are the entries of $\widetilde{B}$, and likewise for B and U.) In particular, the reduction algorithm we have described will produce the factorization (1.2.11) of A with U satisfying condition (1.2.12).

1.3. The algorithm we have described is always feasible (at least theoretically, ignoring ill-conditioning) and it leaves nothing to chance or choice. Thus it, and the resulting decomposition (1.2.11), refines the decomposition (1.1.15) of the "generic" matrix A. It yields a partition of the set of all invertible matrices, i.e., the group GL_n, into a finite number of sets indexed by permutations. In particular, it yields a precise description of the set of "nongeneric" matrices (those for which the factorization (1.1.15) does not exist). Further, the set of matrices of form (1.2.11) for which P is a fixed matrix has a very simple structure. Thus Theorem 1.2.10 suggests GL_n is sort of a "fattened up" version of S_n, the permutation group on n letters. In fact the relation, hinted at in Theorem 1.2.10, between GL_n and S_n is quite intimate, and generalizes to all semisimple Lie groups. This is the first of the connections between Lie groups and finite groups promised at the outset of this section. This linkage was first brought out in the work of Weyl **[Weyl1]**, so S_n and its generalizations are called *Weyl groups*.

1.3.1. To strengthen the reader's belief in the importance of the $S_n - \mathrm{GL}_n$ connection, we point out that the decomposition (1.2.11) has a straightforward and satisfying group-theoretic interpretation. We introduced the group $\overline{\mathscr{U}}$ of unipotent lower triangular matrices. Let $\mathscr{U}$ be the group of unipotent upper triangular matrices.

Then the set of A for which a fixed P and D appear in (1.2.11) is simply a $(\overline{\mathscr{U}}, \mathscr{U})$ double coset. Further, the condition (1.2.12) is simply an irredundancy condition to guarantee that no element of the double coset is written twice. To see this, let $\{u_{ij} : 1 \leq i < j \leq n\}$ be the above diagonal coordinates of a typical element U of $\mathscr{U}$. Then we can check that

$$(1.3.1.1) \qquad \begin{aligned} \mathscr{U} \cap (PD)^{-1}\overline{\mathscr{U}}(PD) &= \mathscr{U} \cap P^{-1}\overline{\mathscr{U}}P \\ &= \{U \in \mathscr{U} : u_{ij} = 0 \text{ if } p(i) > p(j)\}. \end{aligned}$$

This condition is precisely complementary to condition (1.2.12) and only the identity element of $\mathscr{U}$ can satisfy both.

We can carry this further. Let $\mathscr{D}$ be the group of invertible diagonal matrices. Then $\mathscr{D}$ normalizes both $\mathscr{U}$ and $\overline{\mathscr{U}}$, and

$$(1.3.1.2a) \quad \mathscr{B} = \mathscr{D} \cdot \mathscr{U} = \{DU : D \in \mathscr{D}, U \in \mathscr{U}\} = \{UD : D \in \mathscr{D}, u \in \mathscr{U}\}$$

is the group of arbitrary (invertible) upper triangular matrices, and similarly

$$(1.3.1.2b) \qquad \overline{\mathscr{B}} = \mathscr{D}\overline{\mathscr{U}}$$

is the group of lower triangular invertible matrices.

Let W denote the (Weyl) group of permutation matrices. Observe that W normalizes $\mathscr{D}$. Therefore

$$(1.3.1.3) \qquad P\mathscr{D} = \{PD : D \in \mathscr{D}\} = \{DP : D \in \mathscr{D}\} = \mathscr{D}P$$

for any P in W. Therefore, we see that if we only fix P in (1.2.11) and let L, D, and N vary, then we obtain a $(\overline{\mathscr{U}}, \mathscr{B})$, or a $(\overline{\mathscr{B}}, \mathscr{B})$, or a $(\overline{\mathscr{B}}, \mathscr{U})$

double coset. Again, we put D on one side only of P in (1.2.11) to eliminate redundancy.

Thus (1.2.11) implies the double coset decompositions

$$\begin{aligned} \mathrm{GL}_n &= \overline{\mathscr{B}} W \mathscr{B} = \overline{\mathscr{U}} W \mathscr{B} = \overline{\mathscr{B}} W \mathscr{U} \\ &= \bigcup_{P \in W} \overline{\mathscr{B}} P \mathscr{B} = \bigcup_{P \in W} \overline{\mathscr{U}} P \mathscr{B} = \bigcup_{P \in W} \overline{\mathscr{B}} P \mathscr{U} . \end{aligned} \tag{1.3.1.4}$$

We may also observe that $\mathscr{U}$ and $\overline{\mathscr{U}}$, and $\mathscr{B}$ and $\overline{\mathscr{B}}$ are conjugate in GL_n. Explicitly, if

$$w_0 = \begin{bmatrix} 0 & & \cdots & & 1 \\ \vdots & & & 1 & \\ & & & & \\ & & 1 & & \vdots \\ 1 & & \cdots & & 0 \end{bmatrix} \tag{1.3.1.5}$$

is the permutation matrix corresponding to the permutation which exactly reverses order $\{1, 2, \ldots, n\}$, then we see easily that $w_0 = w_0^{-1}$ and

$$\overline{\mathscr{U}} = w_0 \mathscr{U} w_0, \qquad \overline{\mathscr{B}} = w_0 \mathscr{B} w_0. \tag{1.3.1.6}$$

Since $w_0 \in W$, we can combine (1.3.1.4) and (1.3.1.6) to obtain

$$\mathrm{GL}_n = \mathscr{B} W \mathscr{B} = \mathscr{U} W \mathscr{B} = \bigcup_{P \in W} \mathscr{U} P \mathscr{B} . \tag{1.3.1.7}$$

This double coset decomposition of GL_n into $(\mathscr{U}, \mathscr{B})$ double cosets parametrized by the (finite) group W is commonly called the *Bruhat decomposition.* Its analog in a general semisimple or reductive group is a central fact of modern Lie theory. It was described by F. Bruhat **[Bruh]** for several classes of groups. He was motivated by questions in representation theory. It was also observed in several cases by Gelfand and Naimark **[GeNa]**. Its existence in a general semisimple group was established by Harish-Chandra **[HaCh1]**, and it is a central feature of the theory of $(B - N)$-pairs developed by Tits **[Bour, Crtr]**. We will give below some examples of its applications.

1.3.2. Another piece of evidence for the importance of $S_n = W$ in the study of GL_n comes from consideration of the dimensions of the double cosets $\overline{\mathscr{U}} P \mathscr{B}$. Use of the term "generic" for the elements of the identity coset suggests the following:

(1.3.2.1)

(i) The identity coset $\overline{\mathscr{U}} \mathscr{B}$ is an open subvariety of GL_n, of dimension equal to n^2, the same as the dimension of GL_n.

(ii) The other cosets $\overline{\mathscr{U}} P \mathscr{B}$, $P \neq I$, are subvarieties of strictly smaller dimensions.

These statements are true. Further, the codimension of a coset $\overline{\mathscr{U}} P \mathscr{B}$ is a familiar combinatorial function on S_n.

To check statement (1.3.2.1)(i), simply count the number of parameters involved. Elements of the group $\overline{\mathscr{U}}$ have $n(n-1)/2$ lower triangular entries which vary arbitrarily; it has dimension $n(n-1)/2$. Similarly $\mathscr{U}$ has dimension $n(n-1)/2$, and since the diagonal entries of $\mathscr{D}$ are arbitrary subject to being nonzero, it has dimension n. So the dimension of $\overline{\mathscr{U}}\mathscr{B} = \overline{\mathscr{U}}\mathscr{D}\mathscr{U}$ is $2(n(n-1)/2) + n = n^2$.

On the other hand, in describing the coset $\overline{\mathscr{U}}P\mathscr{B}$, we restrict certain of the upper triangular entries u_{ij}, $1 \leq i \leq j \leq n$, of U (as in (1.2.11)) to be zero, according to condition (1.2.12). Condition (1.2.12) says we should set u_{ij}, $i < j$, equal to zero whenever $p(i) > p(j)$, i.e., when p reverses the order of the pair (i, j). Thus the total number of parameters needed to describe the coset $\overline{\mathscr{U}}P\mathscr{B}$ is n^2 minus the number of pairs (i, j) whose order is reversed by p; in other words, the number of pairs reversed by p is the codimension of $\overline{\mathscr{U}}P\mathscr{B}$. But the number of pairs (i, j) whose order is reversed by p is a familiar quantity, usually called the *length* of p **[Bour, Hill]**, and denoted $l(p)$, or also $l(P)$. Here P is, as it has been, the permutation matrix representing p. In summary:

(1.3.2.2) The codimension of the coset $\overline{\mathscr{U}}P\mathscr{B}$ in GL_n is $l(P)$, the length of the permutation associated to P.

Note that $l(P)$ is also the dimension of the subgroup $\mathscr{U} \cap P^{-1}\overline{\mathscr{U}}P$, as described in (1.3.1.1).

1.4. We give here an example of how Theorem 1.2.10 fits into modern mathematics. Consider the set G_k^n of k-dimensional linear subspaces of n-space. The set G_k^n is called a *Grassmann variety* or *Grassmannian*, after Hermann Grassmann (1809–1877), a German Gymnasiumlehrer whose deep geometrical insight was radically under-appreciated during his lifetime. Note that G_1^n is the space of lines in n-space, so is better known as $(n-1)$-dimensional projective space; which of course is the backdrop for classical algebraic geometry. The Grassmannians G_k^n also play a prominent role in classical algebraic geometry **[HoPe]**. But we will discuss here a more recent development.

If Z is a k-dimensional subspace of n-space, and A is in GL_n, then

$$A(Z) = \{A(u) : u \in Z\} \tag{1.4.1}$$

is another k-dimensional subspace. Hence, GL_n acts by permutations on the Grassmannian G_k^n of all k-dimensional spaces. It is an elementary fact in linear algebra that any k-dimensional subspace can be transformed into any other by recipe (1.4.1), for an appropriate choice of A in GL_n. Thus the action of GL_n on G_k^n is transitive, or in other words, G_k^n is a *homogeneous space* or *coset space* for GL_n. Thus, if we choose a base point V_k in G_k^n, we have an identification

$$G_k^n \simeq \mathrm{GL}_n/\mathscr{P}_k, \tag{1.4.2}$$

where $\mathscr{P}_k$ is the stabilizer of V_k—the subgroup of $P \in \mathrm{GL}_n$ such that $P(V_k) = V_k$.

Let us choose for V_k the obvious space of vectors—spanned by the first k standard basis vectors—of the form

$$[x_1, x_2, \dots, x_k, 0, 0, \dots, 0]^t.$$

The stabilizer $\mathscr{P}_k$ of this V_k is easily checked to be the group of matrices of the form

$$\begin{bmatrix} A_1 & X \\ 0 & A_2 \end{bmatrix}, \qquad A_1 \in \mathrm{GL}_k, A_2 \in \mathrm{GL}_{n-k}, X \in M_{k,n-k}. \tag{1.4.3}$$

This group $\mathscr{P}_k$ contains the group $\mathscr{B}$ of upper triangular matrices. It follows from the Bruhat decomposition (1.3.1.7) that GL_n consists of a finite number of $(\mathscr{U}, \mathscr{P}_k)$ double cosets. Under the projection mapping $\pi: \mathrm{GL}_n \to \mathrm{GL}_n/\mathscr{P}_k \simeq G_k^n$, a $(\mathscr{U}, \mathscr{P}_k)$ double coset maps to a $\mathscr{U}$-orbit. Hence we can conclude from (1.3.1.7) that, under the action of $\mathscr{U}$, the Grassmannian G_k^n breaks up into a finite number of orbits.

A finer analysis, amounting to a continuation of the arguments which led to (1.2.11), (1.3.1.7), and (1.3.2.2), yields the following conclusions.

THEOREM 1.4.4. (a) *Set* $W_k = W \cap \mathscr{P}_k$ (*note that via the isomorphism* $W \simeq S_n$ *we have* $W_k \simeq S_k \times S_{n-k}$). *Then the natural inclusion*

$$W/W_k \to \mathscr{U}\backslash\mathrm{GL}_n/\mathscr{P}_k \tag{1.4.5}$$

is a bijection. Hence, under the action of $\mathscr{U}$, *the Grassmannian* G_k^n *consists of the* $\binom{n}{k}$ *orbits*

$$\mathscr{U}\pi(w), \qquad w \in W/W_k. \tag{1.4.6}$$

Here $\pi(w)$ *is the image of* $w \in W$ *in* G_k^n *under the projection* $\pi: \mathrm{GL}_n \to G_k^n \simeq \mathrm{GL}_n/\mathscr{P}_k$.

(b) *Each orbit* $\mathscr{U}\pi(w)$ *is a cell, i.e., may be parametrized in a natural way by a vector space. The dimension of the orbit* $\mathscr{U}\pi(w)$ *is* $l(wW_k)$, *the length, as an element of* S_n, *of the shortest element in the coset* wW_k.[2]

Theorem 1.4.4 has the following consequence. So far the reader may have been thinking of the field of scalars as **R**, the real numbers. We now want them to be **C**, the complex numbers. Then the $\mathscr{U}$-orbits $\mathscr{U}\pi(w)$ are parametrized by complex vector spaces, so their dimensions over **R** are even. Thus in the case of a complex Grassmannian, $G_k^n(\mathbf{C})$, Theorem 1.4.4(b) provides a decomposition into even-dimensional cells. General results in algebraic topology [**Mass**] then guarantee that these cells define a basis for the (rational) homology of $G_k^n(\mathbf{C})$. In other words, Theorem 1.4.4(b) provides direct and detailed information on the topology of the complex Grassmannians; it says we may describe the topology of $G_k^n(\mathbf{C})$ in terms of the combinatorics of S_n.

This is an interesting result in itself, but it acquires still more significance in view of the basic role that Grassmannians and their cohomology play in the theory of vector bundles. We recall **[Ati1, Huse]** that a k-dimensional vector bundle

$$\begin{array}{c} V \\ \downarrow \\ X \end{array} \tag{1.4.7}$$

over a compact Hausdorff space X gives rise to a map (the "classifying map")

$$\gamma_V : X \to G_k^n \tag{1.4.8}$$

for large n. (Observe there is an obvious injection of G_k^n into G_k^{n+1}, so if we have a map (1.4.8) for $n = n_0$, we have such a map for all larger n by composition with these inclusions.) Further, if n is sufficiently large, then the isomorphism class of V is determined by the homotopy class of γ_V.

The pullback map of cohomology

$$\gamma_V^* : H^*(G_k^n) \to H^*(X) \tag{1.4.9}$$

is thus an invariant of the isomorphism class of V. The inverse images under γ_V^* of certain elements of $H^*(G_k^n)$ are the "characteristic classes" (Chern classes, Pontrjagin classes, Todd class, etc.) that figure prominently in the Riemann-Roch formula **[Hirz]**, the index formula **[AtSi, Gilk]** and such matters. These brief indications must suffice for now to suggest how the Bruhat decomposition, which arises in very elementary, classical mathematics, leads directly into sophisticated modern topics.

The discussion given here for G_k^n extends to all "flag varieties", homogeneous spaces of the form $\mathrm{GL}_n/\mathscr{P}$, where $\mathscr{P}$ is any subgroup containing $\mathscr{B}$. Such a variety may be thought of as the set of all nested sequences $\{0\} = V_0 \subseteq V_1 \subseteq V_2 \subseteq \cdots \subseteq V_k$ of subspaces of specified dimensions. There is an analogous theory for all semisimple groups.

1.5. To begin an account of Lie theory with the Bruhat decomposition, though it is consonant with modern views, is unhistorical in the extreme. In particular, the Bruhat decomposition embodies two aspects of Lie theory which were totally lacking at the outset, but which have come to be seen as essential aspects of the theory as it exists today.

First, it is global. The main point of Theorem 1.2.10 is to refine the "generic" analysis leading to the LDU decomposition to an analysis that describes *all* of GL_n. The global aspect of the decomposition (1.2.11) (or (1.3.1.7)) was brought out clearly in the previous remark, when the cosets of the Bruhat decomposition were seen to give rise directly to the homology of the Grassmann varieties. By contrast, the original emphasis of Lie theory was local; only the group in a neighborhood of the identity was considered and calculations were mainly carried out in terms of the "infinitesimal group," now called the Lie algebra. Attention to global features of Lie groups began to be emphasized in the work of Weyl **[Weyl1, Weyl2]**, who also coined the term "Lie algebra."

Second, it is completely algebraic. The alert reader will have noticed that nowhere in the derivation of decomposition (1.2.11) was anything assumed about the nature of the scalars, except that they formed a field, so the usual operations of addition, subtraction, multiplication, and division could be performed. Further, all the groups we dealt with were algebraic groups, i.e., were defined by algebraic equations, and likewise, the double cosets of the Bruhat decomposition, and the corresponding cells in the Grassmann varieties, are all algebraic varieties, and all our constructions were valid over any field.

Appreciation of the importance of the essentially algebraic nature of the theory of semisimple Lie groups did not develop fully until around 1950. The algebraic viewpoint was developed during the 1950s into the theory of algebraic groups by Chevalley, Borel, Tits, and others **[Borel1, 3, 4, Chev1–5, Bour, Tits]**.

This development had at least two major consequences, both of which were important for the theme of this section, the connection between Lie groups and discrete groups.

1.5.1. First, Chevalley **[Chev1]** realized that an algebraic version of Lie theory provided a construction of many simple finite groups. If, in our discussion of Gaussian elimination, we take our scalars to belong to a finite field F_q of q elements, then we are talking about the finite group $\mathrm{GL}_n(F_q)$.[3] This is not a simple group, but it almost is. If we restrict the determinant to be 1, then divide out by the group of scalar matrices, we obtain $\mathrm{PSL}_n(F_q)$, the projective special linear group, which is simple. It had been realized since the 19th century that classical groups (orthogonal, symplectic, unitary, as well as GL_n) have forms over finite fields and that, after elimination of some small abelian groups, these groups give rise to finite simple groups. Also, Dickson **[Dick]** had constructed finite groups corresponding to the exceptional Lie group G_2. However, Chevalley was first to realize the systematic connection between Lie theory, in its incarnation as the theory of algebraic groups, and the construction of finite simple groups. Refinements of his work yielded all infinite series of finite simple groups, leaving out only what are now known to be 26 "sporadic" simple groups. (Some of these, including the largest, the Fischer-Griess "Monster," have very recently been found also to be related to Lie theory in a more subtle way **[FrLM]**.)

1.5.2. Second, the algebraic point of view led to the conception of a very broad class of discrete groups, known as the *arithmetic groups*. Arithmetic groups are important in algebraic geometry and, especially, are an essential part of a vastly generalized formulation of the theory of automorphic forms that developed during the 1950s and 1960s (see §4.2). The precise definition of arithmetic group is technical and not especially enlightening,[4] but the rough idea is that it is a group like GL_n or SL_n, but whose matrices have integer entries. Thus $SL_n(Z)$ is a good example. The point is that arithmetic groups are constructed in a methodical way using algebraic groups.

Consider, by contrast, the simple, abstract condition of being a lattice. A subgroup Γ of a Lie group G is called a *lattice* if

(1.5.2.1)
(i) Γ is discrete.
(ii) The coset space G/Γ carries a finite measure invariant under the permutation action of G.

We observe that if one desires to compare abstract groups with Lie groups, conditions (1.5.2.1) naturally suggest themselves. Imagine we have an abstract group Γ, whose structure we would like to compare with a Lie group G. To make the comparison, we would want to find a homomorphism h of Γ into G. We may as well assume h is an embedding, for G will teach us nothing about $\ker h$. But if h is an injection, we may as well identify Γ with $h(\Gamma)$ and simply consider Γ as a subgroup of G. As a condition of coherence or compatibility between G and Γ, to ensure G is really exercising some control over Γ, we should require Γ to be closed in G. But if Γ is countable, in particular if Γ is finitely generated, this is equivalent to requiring Γ to be discrete. Finally, finiteness of volume of G/Γ is some guarantee that Γ is big enough to "see" all of G.[5] To take a very simple example, any abelian group can be embedded in $\mathbf{R}^n$, provided only that it is torsion-free and of cardinality not greater than the continuum. But, to be embedded discretely, it must be a free group of k generators, with $k \leq n$; and, to be a lattice, it must be free of rank n, i.e., isomorphic to $\mathbf{Z}^n$.

Amazingly, it turns out that the abstract concept "lattice in a Lie group" and the concrete construction "arithmetic group" are, though not identical, very closely related. Thus Borel and Harish-Chandra **[BoHC]** proved that if certain obvious conditions are met, then an arithmetic group is a lattice. For example, $GL_n(\mathbf{Z})$ is not a lattice in $GL_n(\mathbf{R})$, essentially because $GL_1(\mathbf{Z}) = \{\pm 1\}$ is not a lattice in $GL_1(\mathbf{R}) = \mathbf{R}^\times$; but $SL_n(\mathbf{Z})$ is a lattice in $SL_n(\mathbf{R})$. This fact, though nontrivial, is already implicit in the "reduction theory" of Hermite **[Borl3]**, and the proof of Borel and Harish-Chandra may be regarded as a refinement and generalization of this theory.

It is natural to wonder to what extent the converse is true.[6] It clearly is not true, for a very famous reason—the moduli of Riemann surfaces. Every compact Riemann surface X (or surface with a finite number of punctures) can be represented in an essentially unique way (i.e., up to conjugacy of Γ) as a double coset space

(1.5.2.2) $$X \sim SO_2 \backslash SL_2(\mathbf{R})/\Gamma,$$

where

$$SL_2(\mathbf{R}) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} ; a, b, c, d \in \mathbf{R}, ad - bc = 1 \right\}$$

is the group of 2×2, real, determinant one matrices,

$$SO_2 = \left\{ \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} : \theta \in \mathbf{R} \right\}$$

is the subgroup of rotations, and Γ is an appropriate lattice in $SL_2(\mathbf{R})$. But it is well known that the compact Riemann surfaces of a given genus g form a continuous family which fills out a much studied complex manifold of dimension $3g - 3$ (the Teichmuller space) **[Bers]**. As the surface X moves around continuously, so must the corresponding group Γ (as in (1.5.2.2)). In particular, there are uncountably many of them. Since there are only countably many arithmetic groups—for essentially the same reason that there are only countably many algebraic numbers—most lattices in $SL_2(\mathbf{R})$ are not arithmetic.

However, again quite amazingly, this phenomenon of "deformation of lattices" is essentially limited to SL_2. Mostow **[Most1]** showed that if G is a semisimple Lie group, containing no factors (locally) isomorphic to $SL_2(\mathbf{R})$, then a lattice Γ in G is *rigid*, that is, if Γ is another lattice in G and Γ' is sufficiently close (in a fairly straightforward sense) to Γ, then Γ' is actually conjugate to Γ in G.

Thus if we exclude $SL_2(\mathbf{R})$, we can at least wonder if all lattices are perhaps arithmetic. But it is not true. However, Margulis **[Marg1]** (see also **[Zimm1]**) showed that it is very often true. He showed there is a simple condition (that all simple factors have rank at least two) on a semisimple Lie group that guarantees all its lattices are arithmetic. In some sense, "most" Lie groups satisfy Margulis' criterion. For example, any lattice in $SL_n(\mathbf{R})$, $n \geq 3$, is arithmetic. Thus, the theory of algebraic groups has led to a rather deep understanding of the geometric properties of discrete subgroups of Lie groups. I should mention, however, that the question of which semisimple groups contain nonarithmetic lattices is not yet precisely settled. See **[GrPS, Most2]** for examples. This is one problem for the future.

Endnotes. 1. We can in fact formulate precisely the condition that it will work for a given $n \times n$ matrix A. Let A_j be the leading $j \times j$ submatrix of A:

$$A_j = \begin{bmatrix} a_{11}a_{12} & \cdots & a_{1j} \\ a_{21}a_{22} & \cdots & a_{2j} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jj} \end{bmatrix}.$$

Observe that multiplication of A on the left by a lower triangular unipotent matrix does not change $\det A_j$. Thus if $\{d_i\}$ are the diagonal entries of the matrix D in (1.1.15), we have $\det A_j = \det(DU)_j = \prod_{i=1}^{j} d_i$. But in order to carry out the reduction process to achieve the factorization, we need $d_i \neq 0$ for each i. From our formula relating $\det A_j$ to the d_i, we see this condition may be expressed directly in terms of A by the requirement that $\det A_j \neq 0$, $j = 1, 2, \ldots, n$.

2. There is a very beautiful way to count the number of cells of a given dimension. Consider the "Poincaré polynomial"

$$P_{G_k^n}(q) = \sum_{j \geq 0} b_j q^j,$$

where b_j is the number of cells (i.e., $\mathscr{U}$-orbits) of dimension j. Consider also the Poincaré polynomial

$$P_W(q) = P_{S_n}(q) = \sum_{w \in W} q^{l(w)}$$

whose coefficients count the number of elements of W of given length. It turns out that P_{W_k} divides P_W, and

$$P_{G_k^n} = P_W / P_{W_k} = P_{S_n} / (P_{S_k} \cdot P_{S_{n-k}}).$$

Explicitly, one has

$$P_{S_n} = \prod_{i=1}^{n} \frac{q^i - 1}{q - 1}.$$

Hence

$$P_{G_k^n} = \prod_{i=1}^{k} \frac{q^{n-k+i} - 1)}{q^i - 1}.$$

The $P_{G_k^n}$ are well known in combinatorics as the "Gaussian polynomials" **[Proc1, Zeil3]**.

3. In this situation, the double coset $\mathscr{U} P \mathscr{B}$ of (1.3.17) can be seen to have order $q^{n(n-1)/2}(q-1)^n q^{l(P)}$, where $l(P)$ is the length of the permutation P. Summing over P gives the formula

$$^{\#}(\mathrm{GL}_n(F_q)) = q^m (q-1)^n \sum_{p \in S^n} q^{l(P)} = q^m (q-1)^n P_{S_n}(q),$$

where $m = n(n-1)/2$ and P_{S_n} is the Poincaré polynomial of S_n. Comparison with the easily derived formulas

$$^{\#}(\mathrm{GL}_n(F_q)) = \prod_{i=0}^{n-1} (q^n - q^i)$$

gives a formula for $P_{S_n}(q)$. A similar method applies to the Poincaré polynomial for Grassmannians (cf. note 2).

This is a very modest example of the transferral of information from characteristic p to characteristic zero—a method which, thanks to recent developments in algebraic geometry, has become extremely powerful. Some spectacular examples of it are the Deligne-Lusztig construction of representations of finite Chevalley groups **[DeLu]**, and the Beilinson-Bernstein and Brylinski-Kashiwara proofs of the Kazhdan-Lusztig conjectures **[BeBe, BrKa]**.

4. For the 'satiably curious, here it is: A (linear) algebraic subgroup of $\mathrm{GL}_n(\mathbf{C})$ is a subgroup which is also an algebraic subvariety, i.e., the set of zeros of a collection of polynomials in X_{ij}, the coordinate functions on M_n, and in $\det^{-1}$, the reciprocal of the determinant function. Thus $\mathrm{SL}_n(\mathbf{C})$, the special linear group, is defined by the equation $\det g = 1$; and $\mathrm{O}_n(\mathbf{C})$, the complex orthogonal group, is defined by the equation $g^t g = 1$, which can be regarded as the collection of n^2 scalar equations

$$\sum_{j=1}^{n} g_{ji} g_{jk} = \delta_{ik} \quad \text{for } 1 \leq i, k \leq n.$$

Let $G \subseteq \mathrm{GL}_n(\mathbf{C})$ be an algebraic subgroup, and let I_G be the ideal of polynomials vanishing on G. If we can find a generating set $\{P_j\}_{j=1}^{m}$ of polynomials for I_G such that the coefficients of the P_j are real, we say G is *defined over* $\mathbf{R}$. If we can find such P_j with coefficients in $\mathbf{Q}$, we say G is *defined over* $\mathbf{Q}$. If G is defined over $\mathbf{R}$, then

$$G_{\mathbf{R}} = G \cap \mathrm{GL}_n(\mathbf{R})$$

is called the *real points* of G. Similarly for $\mathbf{Q}$.

If G is a linear algebraic group defined over $\mathbf{Q}$, then $G \cap \mathrm{GL}_n(\mathbf{Z}) = G_{\mathbf{R}} \cap \mathrm{GL}_n(\mathbf{Z})$ is called the subgroup of *integral points of* G. Two subgroups Γ_1, Γ_2 of a group G are called *commensurable* if $\Gamma_1 \cap \Gamma_2$ has finite index in both Γ_1 and Γ_2.

Let G_0 be a Lie group, and $\Gamma \subseteq G_0$ a discrete subgroup. We say Γ is *arithmetic* if there exist

(i) a linear algebraic group $G \subseteq \mathrm{GL}_n(\mathbf{C})$ defined over $\mathbf{Q}$, and
(ii) a homomorphism $\Psi: G_0 \to G_{\mathbf{R}}$, such that
(iii) $\ker \Psi$ is compact,
(iv) $\mathrm{im}\, \Psi$ is normal in $G_{\mathbf{R}}$,
(v) $G_{\mathbf{R}}/\mathrm{im}\, \Psi$ is compact, and, most importantly,
(vi) $\Psi^{-1}(G \cap \mathrm{GL}_n(\mathbf{Z}))/\ker \Psi$ is commensurable with $(\Gamma \cdot \ker \Psi)/\ker \Psi$.

5. It might be thought one should demand that G/Γ be compact; but this would exclude important examples including $\mathrm{SL}_n(\mathbf{Z})$. The relaxation of compactness to finiteness of covolume has been very fruitful.

6. A precise question of this nature was formulated by Selberg **[Selb1]**.

2. An outline of Lie theory.

2.1. The glue that binds Lie theory together is the notion of a one-parameter group and its infinitesimal generator. For expository purposes, we will consider the one-parameter group first, although this is a revisionist way of proceeding. This discussion can be found in many texts (e.g., **[Ster, Gilm, Pont, Hoch, Vara]**, etc.). But it is so basic, it seems necessary to include it.

Let M be a manifold. (You can just think of an open set in $\mathbf{R}^n$ if you wish.) The set $\mathrm{Diff}(M)$ of diffeomorphisms (smooth, smoothly invertible